DOCUMENT RESUME

ED 249 230 TM 840 457

TITLE	Student Achievement in Department of Defense Dependents Schools: 1983-84 Annual Test Report.
INSTITUTION	Dependents Schools (DOD), Washington, D.C.
REPORT NO	DoDDS-84-C-0002
PUB DATE	Mar 84
NOTE	29p.
PUB TYPE	Statistical Data (110) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Academic Achievement; Achievement Tests; *Basic Skills; *Educational Assessment; Elementary Secondary Education; Grade 1; Language Arts; Mathematics Achievement; Reading Achievement; *School Readiness; School Readiness Tests; Scores; Testing Programs; Test Norms; *Test Results
IDENTIFIERS	*Dependents Schools; Dependents Schools Systemwide Testing Program; Metropolitan Readiness Tests

ABSTRACT

As part of the Department of Defense Dependents Schools (DoDDS) annual assessment of achievement, the Metropolitan Readiness Tests (MRT) were administered to approximately 14,768 first-grade pupils in the fall of 1983. The mean scaled scores of the DoDDS first graders clustered around the national mean of 150. The DoDDS mean scores in the various skill areas ranged from a high of 7.5 points above the national mean in auditory skills to a low of 1.1 points below the national mean in language. The DoDDS basic skills achievement testing program was administered to 47,861 students in all five DoDDS regions. Known as the Systemwide Testing Program, it measured the performance of students enrolled in grades 4, 6, 7, 9, and 11 in the basic skill areas of reading, language arts, and mathematics. DoDDS students on the average performed better than their stateside counterparts. Overall performance on the tests indicated high achievement for grades 6, 7, 9, and 11, with grade 7 evidencing the highest performance. In the content areas, language arts was observed as the strongest area, followed by reading and mathematics, respectively. By contrast, a notable weakness was evidenced in grade 4 mathematics. (BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

STUDENT ACHIEVEMENT IN DEPARTMENT OF DEFENSE DEPENDENTS SCHOOLS

ED249230

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

B. Stephens

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

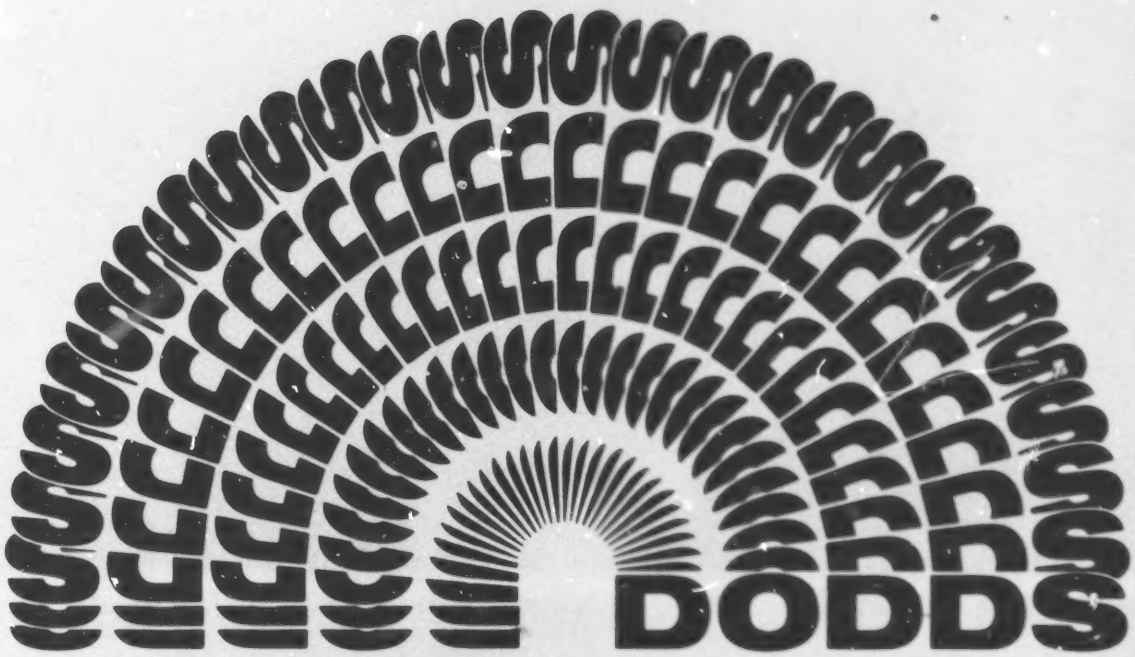
U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

1983-84 ANNUAL TEST REPORT

TM 840 457



CTB/McGraw-Hill

ANNUAL
TEST REPORT

School Year 1983-84

Beth Stephens, Ph.D.
Director
Department of Defense Dependents Schools

TABLE OF CONTENTS

Page

SUMMARY OF FINDINGS v

INTRODUCTION vi

PART I - FIRST GRADE SCHOOL READINESS TESTING PROGRAM
- Introduction 1
- Description of the Tests 1
- Methodology 1
- Results 2

PART II - SYSTEMWIDE TESTING PROGRAM
- Introduction 7
- Description of the Tests 7
- Methodology 9
- Results 10
- Demographic Information 17

APPENDIX A 19

APPENDIX B 20

INTRODUCTION

The annual assessment of achievement of the Department of Defense Dependents Schools (DoDDS) students in selected elementary and secondary school grades was conducted in accordance with Section 1405 (a) and (b), Title XIV, Defense Dependents' Education Act of 1978, PL 95-561, and DS Regulation 2000.6.

Policy

A. Annual assessment of achievement in the subject areas of reading, language arts, and mathematics shall be administered to all pupils enrolled in specified primary, elementary, and secondary grades.

B. An annual school readiness assessment shall be administered to all pupils entering first grade.

C. The individual pupil results of the assessments shall be reported to the pupils, their parents, and their classroom teachers. These results shall become part of the pupil's academic record.

D. Within 90 days of completion of the assessments, summary results of group performance for each of the annual assessments shall be made available to the professional educators of the dependents' schools system and to Members of Congress.

DoDDS grade 1 students were administered the Metropolitan Readiness Tests (MRT), published by the Psychological Corporation. Students enrolled in grades 4, 6, 7, 9, and 11 were administered the Systemwide Testing Program (STP), which consists of reading, language arts, and mathematics basic skill measures developed for DoDDS by CTB/McGraw-Hill Publishing Company.

The assessment administration occurred during the fall of 1983. Students in grade 1 were tested with the MRT the first 2 weeks of school, while students in grades 4, 6, 7, 9, and 11 were administered the STP near the end of September. The 1983-84 assessment involved approximately 63,000 dependents schools' students spread across the five DoDDS regions.

Results from both programs of the assessment were provided for each student, classroom, school, region, and for the DoDDS system as a whole.

This document reports the results to the Congress, school personnel, parents, and other interested citizens. The results are presented in two parts: Part I - First Grade School Readiness Testing Program and Part II - Systemwide Testing Program.

SUMMARY OF FINDINGS

Metropolitan Readiness Tests

The Metropolitan Readiness Tests (MRT) were administered to approximately 14,768 first-grade pupils in the fall of 1983. The MRT is designed to provide a skill-based assessment of those enabling skills that are important for early school learning in reading, mathematics, and language development.

The mean scaled scores of the Department of Defense Dependents Schools (DoDDS) first graders, as a group, cluster around the national mean of 150. The DoDDS mean scores in the various skill areas ranged from a high of 7.5 points above the national mean in auditory skills to a low of 1.1 points below the national mean in language. It was found that students who were reported to be proficient in English scored substantially higher than students who were not proficient in English.

Systemwide Testing Program

The DoDDS basic skills achievement testing program was administered during the last 2 weeks of September 1983 to 47,861 students in all five DoDDS regions. Known as the Systemwide Testing Program, it measured the performance of students enrolled in grades 4, 6, 7, 9, and 11 in the basic skill areas of reading, language arts, and mathematics.

DoDDS students on the average perform better than their stateside counterparts. Overall performance on the tests indicates high achievement for grades 6, 7, 9, and 11, with grade 7 evidencing the highest performance. In the content areas, language arts is observed as the strongest area, followed by reading and mathematics, respectively. By contrast, a notable weakness is evidenced in grade 4 mathematics. Performance in both multiplication and division computation substantially lowered the fourth-grade scores.

PART I
FIRST GRADE READINESS TESTING PROGRAM

INTRODUCTION

During the week of September 12-16 of school year 1983-84, all entering first-grade pupils in the Department of Defense Dependents Schools (DoDDS) were administered the Metropolitan Readiness Tests (MRT), Level II, Form P. This report describes the testing procedures and presents a summary of the results.

DESCRIPTION OF THE TESTS

The MRT, published by the Psychological Corporation, is a widely-used standardized battery of tests designed to measure the school readiness skills associated with beginning reading and mathematics instruction. The 1976 edition of the MRT was normed in a nation-wide standardization project involving over 100,000 students in 322 school districts representing a cross-section of schools in the United States. The standardization process offers the capability of comparing DoDDS scores with the scores obtained by a representative national sample of students.

Four broad skill areas are measured, each of which contains two subtests as specified below:

SKILL AREA	SUBTESTS
Auditory	Beginning Consonants Sound-Letter Correspondence
Visual	Visual Matching Finding Patterns
Language	School Language Listening
Quantitative	Quantitative Concepts Quantitative Operations

Scores are generated for each of the skill areas. In addition, the auditory, visual, and language subtests are combined to produce a prereading composite score.

METHODOLOGY

A total of 14,768 entering first-grade pupils were administered the MRT battery. Table 1 summarizes the DoDDS regional participation in the assessment project.

Classroom teachers administered the test battery using the standardized test administration procedures outlined in the teachers' manual. Fall norms for large city school systems were used to generate descriptive data, including group means, standard deviations, percentile rank scores, and stanines. Standardized scaled scores were produced to compare performance across skill areas.

Table 1

Total Number of Students Administered the MRT by DoDDS Region

Atlantic	Germany	Mediterranean	Pacific	Panama	DoDDS Total
1311	8885	1235	2676	661	14,768

RESULTS

In Table 2 are the mean percentile rank scores and the mean scaled scores for each of the skill area composites for DoDDS. Table 3 presents the regional scores. The mean scaled scores allow for comparisons across skill areas. These data suggest a somewhat even performance across the skill areas, with a scaled score range from a low of 148.9 (language) to a high of 157.2 (auditory). Performance was strongest in the auditory skill area, which assesses perceptual recognition of beginning letter consonants and sound-to-letter correspondence. Performance was lowest on the language skill subtests, which test for listening comprehension, use of grammatical structures, and comprehension of school-related language concepts.

Seventy-four percent of the students were reported to have at least average proficiency in English. (See Table 4.) Those students averaged 155.7 scaled score points in the quantitative area and 157.9 in the prereading area. Students below average in English proficiency scored 23.8 points lower in the quantitative area (131.9) and 23.7 points lower in the prereading area (134.2). A substantial discrepancy in scores exists also between pupils who attended kindergarten and pupils who did not attend kindergarten. (See Table 5.) However, caution must be taken when interpreting the data, due to the large percentage of pupils for whom there were no responses to these questions.

Table 2

DoDDS Scores by Skill Area

Skill Area	Percentile Rank of Mean Raw Score	Mean Scaled Score	
Auditory Composite	65	157.2	(25.6)
Visual Composite	50	152.8	(28.5)
Language Composite	35	148.9	(29.6)
Quantitative Composite	40	150.3	(26.4)
Prereading Composite	60	152.5	(25.1)

150 = National Mean Scaled Score ( ) = Standard Deviation

Table 3

Regional Mean Scores by Skill Area

Skill Area		Atlantic	Germany	Mediterranean	Pacific	Panama
Auditory	$\bar{X}$	158.5	155.8	153.6	161.4	162.8
	SD	26.0	25.6	26.6	23.9	25.7
	PR	65	60	55	75	75
Visual	$\bar{X}$	152.1	151.6	149.9	158.2	154.3
	SD	28.4	29.0	27.2	27.4	27.5
	PR	45	45	40	65	50
Language	$\bar{X}$	154.7	149.7	146.3	145.1	146.1
	SD	29.8	29.1	29.4	30.2	31.1
	PR	50	35	30	25	25
Quantitative	$\bar{X}$	153.8	149.2	146.1	154.7	148.1
	SD	26.7	26.4	25.9	25.7	27.6
	PR	50	40	30	55	35
Prereading	$\bar{X}$	154.2	151.5	149.0	155.7	154.9
	SD	25.2	25.7	24.5	23.0	24.4
	PR	60	55	50	65	65

$\bar{X}$ = Mean Scaled Score
SD = Standard Deviation
PR = Percentile Rank
150 = National Mean Scaled Score

Table 4

Quantitative and Prereading Composite Scaled Scores By Response to Language-Related Demographic Questions

Question/Response	Quantitative Mean Scaled Score	Prereading Mean Scaled Score	Percent Responding
How proficient is the child in English?			
Below Average	131.9	134.2	7
Average	149.2	151.1	45
Above Average	160.6	163.1	12
High	157.3	159.6	17
No Response	147.0	149.1	19
Does the child experience a home language other than English?			
Yes	146.7	149.0	19
No	151.8	154.0	66
No Response	147.8	150.2	15

Table 5

Quantitative and Prereading Mean Scaled Scores By Response to Kindergarten Attendance Questions

Response	Quantitative Mean Scaled Score	Prereading Mean Scaled Score	Percent Responding
Attended Overseas in DoDDS School	150.1	152.4	64
Attended Overseas in Non-DoDDS School	151.3	152.2	.2
Attended Stateside	154.7	156.6	13
Did Not Attend	137.5	139.2	1
No Information	149.0	149.3	3
No Response	148.6	151.0	17

The mean scaled scores of DoDDS first graders as a group cluster around the national mean scaled score of 150. This is substantiated by the regional data in Table 3 and the observed percent of scores within each stanine in Table 6. Table 6 illustrates the clustering of performance within the middle three stanines (4-6), which exceeds the expected stanine distribution from the test standardization. Fewer pupils than expected are scoring in the lower and upper stanine ranges on the prereading and quantitative composites.

Substantial score increases from school year 1982-83 to school year 1983-84 are evident in all skill areas. (See Figure 1.) Increases of 10 percentage points are found in the auditory, visual, language, and prereading areas, whereas the quantitative increase was 5 percentage points.

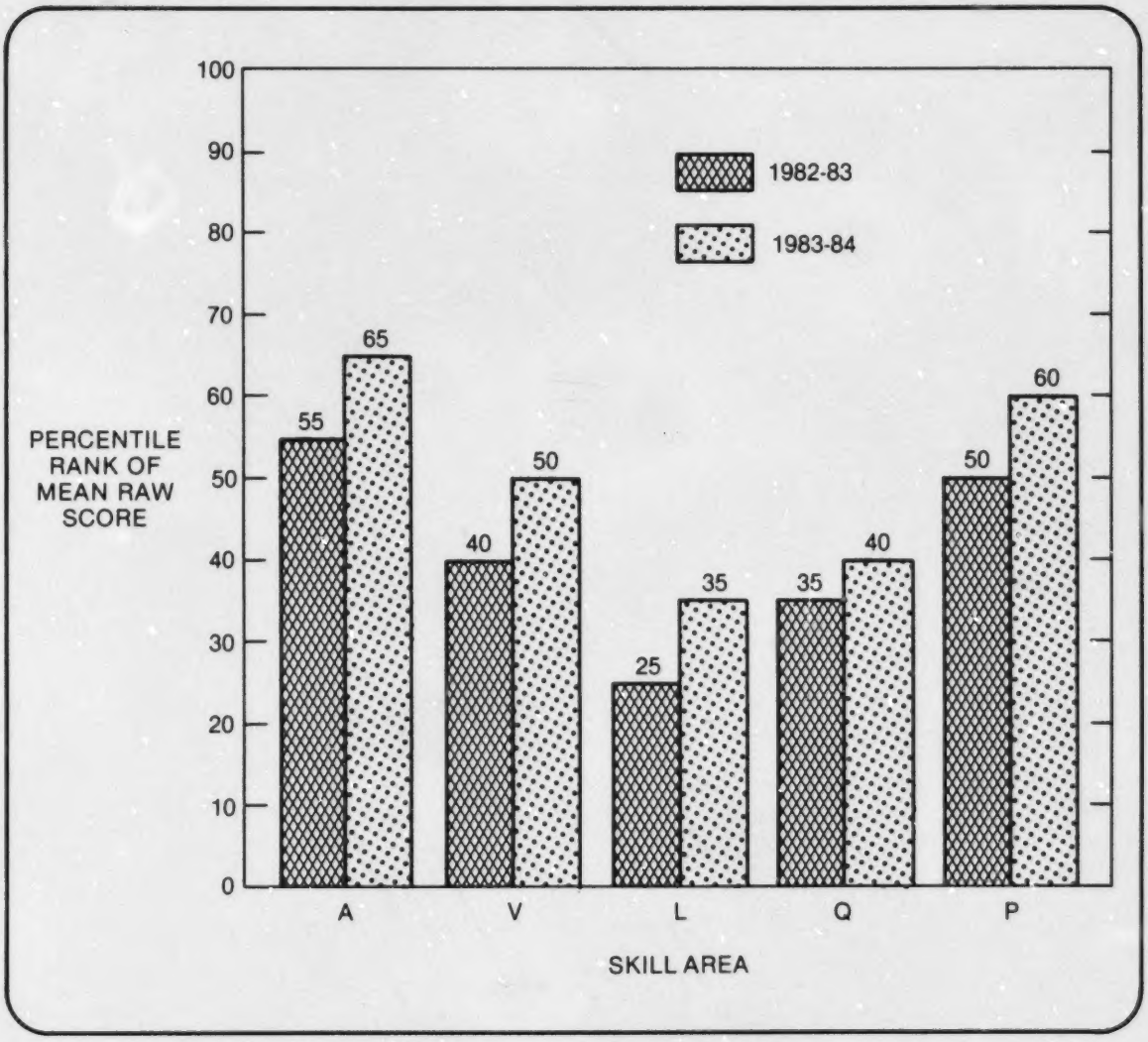
Table 6

Observed and Expected Percent of DoDDS Pupils Scoring Within Each Stanine

Stanine	Expected Percent	Observed Percent of Scores Within Each Stanine				
		Auditory	Visual	Language	Prereading	Quantitative
9	4.0	5.4	3.0	5.8	3.1	1.8
8	7.0	9.8	5.3	6.9	4.0	3.0
7	12.0	10.0	14.8	9.4	11.8	10.5
6	17.0	23.8	15.1	11.1	19.9	15.4
5	20.0	23.5	24.2	23.7	27.5	24.4
4	17.0	15.0	18.0	20.5	19.2	26.8
3	12.0	8.5	10.9	12.9	9.3	12.1
2	7.0	2.7	6.3	6.9	3.6	3.9
1	4.0	1.4	2.4	2.8	1.6	2.1
High	23	25.2	23.1	22.1	19.0	15.3
Average	54	62.2	57.3	55.3	66.5	66.6
Low	23	12.6	19.6	22.6	14.5	18.1

Figure 1

DoDDS Percentile Rank Scores by Skill Area:
SY 1982-83 Versus SY 1983-84



PART II
SYSTEMWIDE TESTING PROGRAM

INTRODUCTION

The DoDDS Systemwide Testing Program (STP) measures the status of basic skills achievement. Students in grades 4, 6, 7, 9, and 11 were administered the STP reading, language arts, and mathematics tests during the week of September 26, 1983.

DESCRIPTION OF THE TESTS

The STP tests are standardized measures of basic skills achievement for reading, language arts, and mathematics developed by CTB/McGraw Hill Publishing Company working in conjunction with DoDDS evaluation and curriculum personnel. CTB/McGraw-Hill designed the test instruments to collect student performance data indicating learned attainment relative to DoDDS educational objectives and to reflect any differences in performance occurring across grade levels. The 25 objectives included in the tests represent the wide array of skills taught in word knowledge, reading comprehension, language usage, and mathematics application, and represent information students are expected to know at the tested grade levels.

The Reading Test assessed student performance in three domains: word attack, vocabulary, and reading comprehension. Word attack was assessed at the fourth grade level only. Vocabulary and reading comprehension were assessed across the five grade levels.

Likewise, student performance in language arts was assessed in three domains: spelling, language mechanics, and language expression.

The mathematics test assessed student performance in two domains: math computation and mathematical concepts and applications.

Locational skills, which might be considered a part of either the reading or language arts curricula, was treated as a separate domain and tested at four grade levels: 6, 7, 9, and 11.

The number of items in these domains ranged from 4 to 23.

Table 7 presents the content coverage of the tests for each grade. A description of the tests' category objectives is included as Appendix A, and the test development procedures are included as Appendix B.

Table 7

STP Content Coverage

Subtests	Skill Areas	Number of Items at Each Grade Level				
		4	6	7	9	11
Word Attack						
	Letter-Sound Correspondence	6				
	Structural Analysis	9				
	Phonic Analysis	5				
Vocabulary						
	Word Meaning	23	23	23	23	23
Reading Comprehension						
	Literal Comprehension	12	9	7	5	5
	Literary Works	7	8	8	6	9
	Higher-Level Comprehension	4	6	8	12	9
Spelling						
	Spelling	20	20	20	20	20
Language Mechanics						
	Capitalization	10	5			
	Punctuation	6	9	11	10	10
	Editing Skills	4	6	9	10	10
Language Expression						
	Usage	9	8	5	4	5
	Sentence Development	5	6	7	7	4
	Paragraph Development	6	9	11	12	13
Math Computation						
	Addition	5	5	5	5	5
	Subtraction	5	5	5	5	5
	Multiplication	5	5	5	5	5
	Division	5	5	5	5	5
Math Concepts and Applications						
	Number Awareness	8	9	9	6	7
	Problem Solving	8	9	9	12	11
	Measurement/Geometry	4	4	4	4	4
Locational Skills						
	Book Parts		12	4	9	9
	Dictionary Skills		4	4	5	
	Library Skills		4	8	6	11
	Consumer Skills			4		

METHODOLOGY

The DoDDS regional evaluation coordinators, building coordinators, and examiners were responsible to the regional directors for the proper administration and security of DoDDS STP materials.

In preparation for the test administration, the regional evaluation coordinators were provided in-service training to assist them in conducting workshops with building coordinators.

After attending building coordinators' training sessions, the building coordinator scheduled and conducted training for the classroom examiners. The building coordinator used the Building Coordinator's Manual and Examiner's Manual with other appropriate materials to conduct the training session.

Tests were administered the last week of September 1983. A total of 47,861 students across the five DoDDS regions participated in the assessment. The Germany Region had the largest percentage of students participating in the program, followed by the Pacific, the Atlantic, the Mediterranean, and Panama regions. Table 8 displays the number and percentage of participating students by grade and region.

Practice tests were available and administered in grades 4 and 6. At the completion of testing, each individual school within the regions assembled and packaged answer sheets and shipped them to the contractor, CTB/McGraw-Hill Scoring Center, Durham, North Carolina. The answer sheets were scored, reports were generated, and the results returned to the DoD schools.

Table 8

Number and Percent of Students in Each Grade Level and Region Participating in DoDDS STP

Fall 1983

Grade	Atlantic	Germany	Med.	Pacific	Panama	Total	% Grade
4	1318	7057	1172	2180	553	12,280	25.6
6	1248	6206	1177	2013	546	11,190	23.4
7	1242	5966	1109	1888	495	10,700	22.4
9	1002	4506	864	1375	492	8,239	17.2
11	725	2930	531	899	367	5,452	11.4
Total N	5,535	26,665	4,853	8,355	2,453	47,861	
Regional Percentage	11.6	55.7	10.1	17.5	5.1		100.0

STP provides two types of scores to help assess the instructional needs of students: norm-referenced and criterion-referenced scores. The norm-referenced score information answers the question, "How do DoDDS students compare with the national norm?" The criterion-referenced information answers the question, "How well are DoDDS students mastering the curriculum objectives?" In providing answers to these questions, the test results are reported in a variety of formats designed to best meet the needs of the user. For example, the classroom reports are organized in a way that allows the teacher to focus on and utilize relevant information about individual students and the class as a whole. The data are reported on other group reports (school, region, and system-wide) in such a way that contrast can be made between the extent to which different educational objectives are being attained and what most needs to be accomplished.

RESULTS

The scale score is the basic score for DoDDS STP. It is used primarily to provide a basis for deriving the percentile scores which are used to describe test performance.

Table 9 and Figures 2, 3, and 4 present the mean scaled score comparisons between DoDDS students and the national norming sample.

Table 9

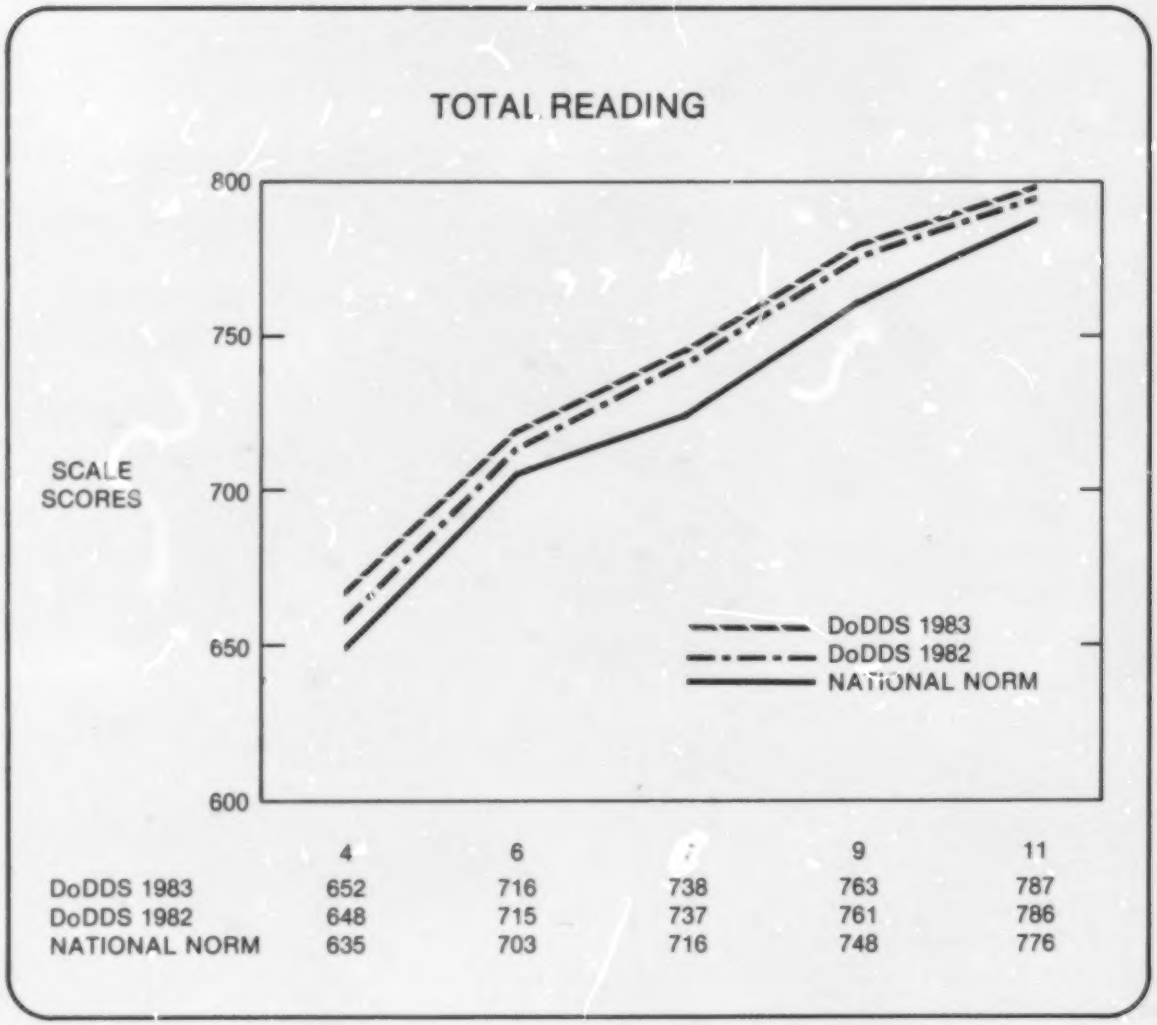
Mean Scaled Score Comparison Between DoDDS Students and National Norming Sample

Grades	4	6	7	9	11
Total Reading					
DoDDS	652	716	738	763	787
National Norm	635	703	710	748	776
S.D.	58.8	51.8	43.4	44.1	39.4
Total Language					
DoDDS	650	696	716	737	750
National Norm	636	684	692	713	732
S.D.	42.2	39.2	44.2	48.9	43.2
Total Mathematics					
DoDDS	646	697	714	733	742
National Norm	648	692	704	723	735
S.D.	39.7	23.8	24.1	23.1	23.3

S.D. = Standard Deviation

Figure 2

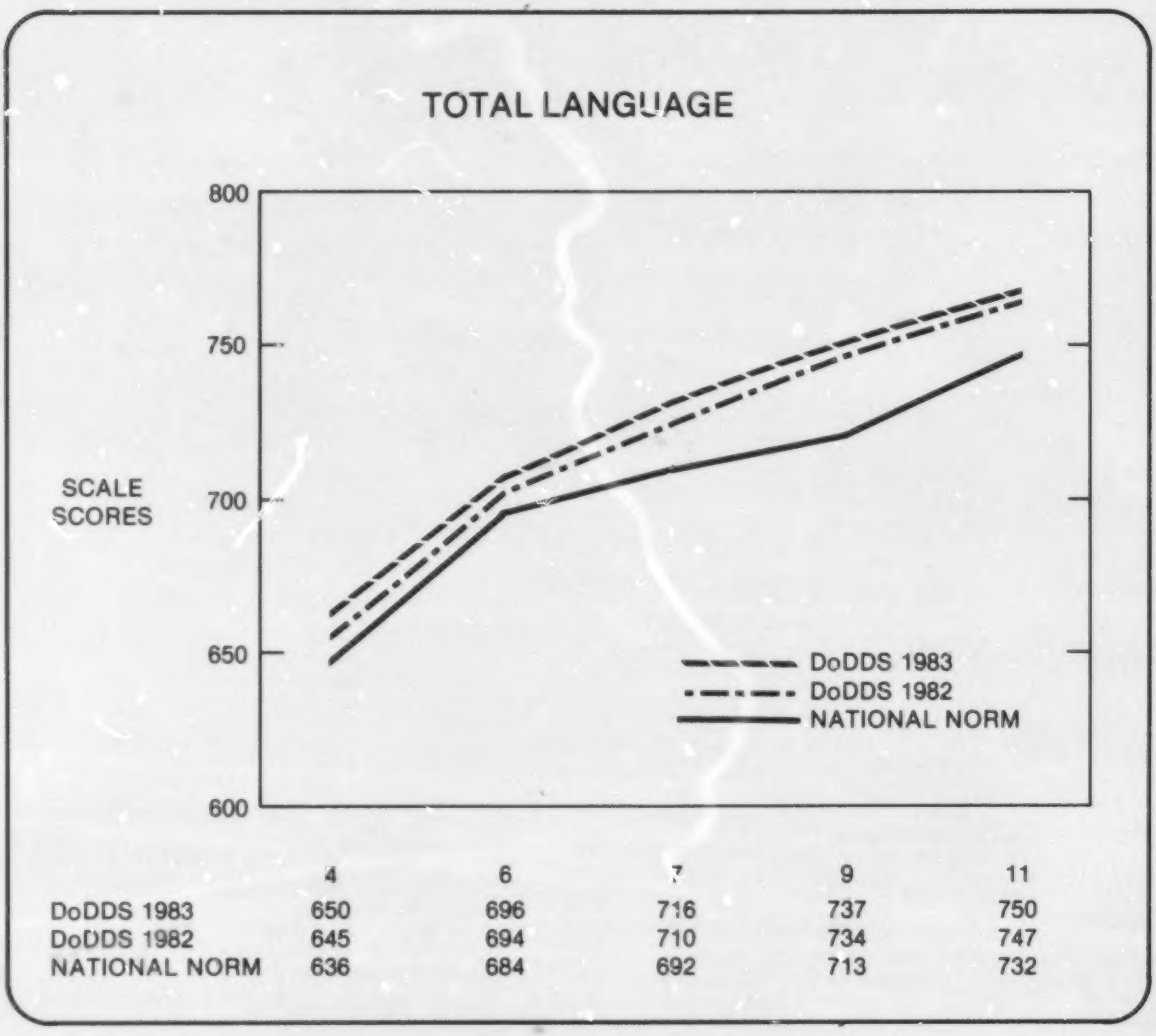
DoDDS Systemwide Testing Program
Comparison of Scaled Scores
1982 and 1983



	4	6	7	9	11
DoDDS 1983	652	716	738	763	787
DoDDS 1982	648	715	737	761	786
NATIONAL NORM	635	703	716	748	776

Figure 3

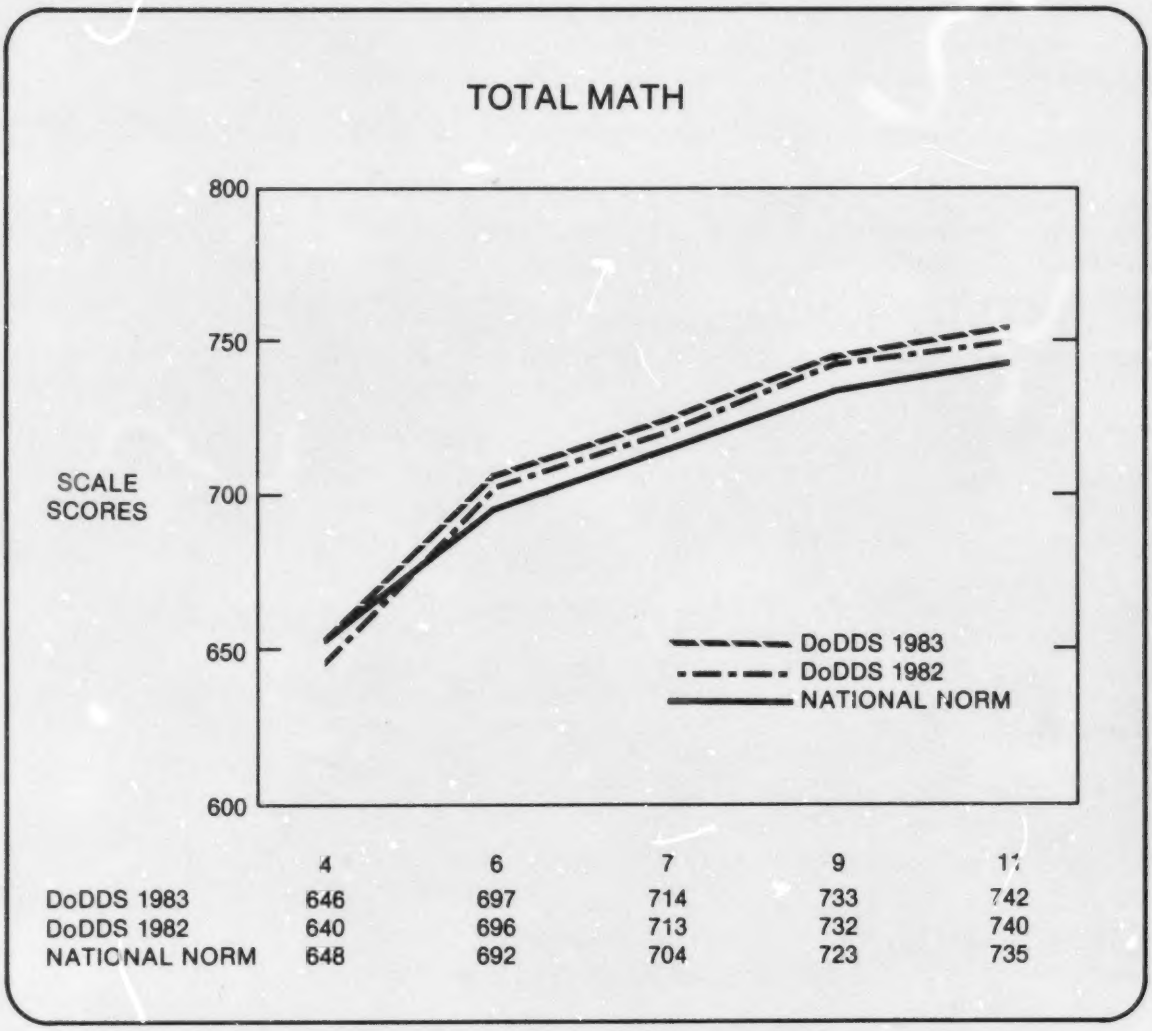
DoDDS Systemwide Testing Program
Comparison of Scaled Scores
1982 and 1983



	4	6	7	9	11
DoDDS 1983	650	696	716	737	750
DoDDS 1982	645	694	710	734	747
NATIONAL NORM	636	684	692	713	732

Figure 4

DoDDS Systemwide Testing Program
Comparison of Scaled Scores
1982 and 1983



	4	6	7	9	11
DoDDS 1983	646	697	714	733	742
DoDDS 1982	640	696	713	732	740
NATIONAL NORM	648	692	704	723	735

STP scores are typically reported by percentile rankings. Percentile rank scores range from 1 through 99. A percentile rank may be interpreted as the percentage of students whose scores fall below a particular group's scale score. For example, if a group's mean scale score converts to a percentile rank of 71, this indicates that the group scored higher than approximately 71 percent of the norming population. In Table 11 are the DoDDS mean percentile scores by grade. A review of the table shows that at grades 6, 7, 9, and 11, DoDDS students' overall performance exceeds the national norm in all content areas. Students in grade 4 attain a high level of performance in reading and language arts, but evidence weakness in mathematics.

The mean percentile scores for Total Reading, Total Language Arts, and Total Mathematics are displayed in Table 10. Overall performance on the tests indicates high achievement for grades 6, 7, 9, and 11, with grade 7 evidencing the highest performance. In the content areas, language arts is observed as the strongest, followed by reading and mathematics, respectively. By contrast, a notable weakness is evidenced by grade 4 mathematics as noted in Table 11.

Performance in both multiplication and division computation substantially lowered the fourth grade scores. When scores from 1982 and 1983 are compared, the data show increases in achievement across the grades, with the most significant change noted for grade 4.

The regional performance and the systemwide results for SY 1982-83 and SY 1983-84 are presented in Table 12. These data allow for the comparison of student performance. Analysis of these data reveal that DoDDS students continue to achieve at higher levels of performance than their U.S. counterparts.

Favorable increase in achievement is observed across the grades with strongest performance noted for language arts followed by reading and mathematics, respectively. The data also show that fourth graders still evidence low performance in mathematics, achieving scores which fall below the national norm. However, when the data are examined for significant improvements, the greatest increase in achievement is observed for grade 4, especially in the area of mathematics.

Table 10

DoDDS Mean Percentile Scores* by Grade

Grade	Total Reading	Total Language Arts	Total Mathematics
4	61	62	48
6	59	61	58
7	65	73	71
9	62	70	67
11	59	67	68

National Norm = 50th Percentile

*Percentiles Based on the Mean Scale Score

Table 11

Subtest and Total Test Mean Percentile Scores by Grade

READING			
Grade	Vocabulary	Comprehension	Total
4	56	65	61
6	59	59	59
7	62	66	65
9	60	63	62
11	53	65	59

LANGUAGE ARTS			
Grade	Mechanics	Expression	Total
4	53	71	62
6	57	58	61
7	65	72	73
9	67	72	70
11	62	66	67

MATHEMATICS			
Grade	Computation	Concepts and Application	Total
4	30	64	48
6	58	55	59
7	61	73	71
9	62	64	67
11	80	54	68

Table 12

Mean Percentile Scores on DoDDS Systemwide Testing Program by Grade and Region

READING

Grade	Atlantic 82	Atlantic 83	Germany 82	Germany 83	Mediterranean 82	Mediterranean 83	Pacific 82	Pacific 83	Panama 82	Panama 83	System 82	System 83
4	65	68	60	60	61	63	59	61	63	62	59	61
6	67	65	59	58	62	63	61	59	60	62	59	59
7	69	68	64	64	66	68	66	67	64	68	64	65
9	68	68	58	59	63	63	63	63	57	63	60	62
11	64	68	57	57	59	61	58	59	58	61	58	59

LANGUAGE ARTS

Grade	Atlantic 82	Atlantic 83	Germany 82	Germany 83	Mediterranean 82	Mediterranean 83	Pacific 82	Pacific 83	Panama 82	Panama 83	System 82	System 83
4	61	64	60	62	62	62	61	64	64	67	57	62
6	62	63	60	58	65	65	63	65	62	68	59	61
7	73	75	66	70	72	76	71	74	65	75	67	73
9	74	76	65	67	72	73	74	74	60	72	67	70
11	69	73	62	64	66	72	65	68	61	67	64	67

MATHEMATICS

Grade	Atlantic 82	Atlantic 83	Germany 82	Germany 83	Mediterranean 82	Mediterranean 83	Pacific 82	Pacific 83	Panama 82	Panama 83	System 82	System 83
4	45	52	44	46	38	41	50	56	41	57	43	48
6	64	62	57	54	62	64	62	62	57	62	59	59
7	74	72	66	69	72	79	74	72	69	79	69	71
9	75	75	63	63	67	71	71	69	61	69	65	67
11	70	76	61	64	68	72	68	72	64	64	64	68

The following observations about changes in achievement are also evident from an examination of the data.

Reading

- Reading test scores improved in four of the five grades tested; namely, grades 4, 7, 9, and 11.
- Highest scores were achieved in the Atlantic Region; however, two grades show a loss.
- The Panama Region shows significant reading score increase at grade 6.
- The Panama Region, followed by the Mediterranean, shows the most consistent increase in performance.

Language Arts

- All regions show language arts score improvement at grade 11.
- The Panama Region shows the highest performance of the regions for the lower grades, and the Atlantic Region scores are the highest at the upper grades.
- The Panama Region shows dramatic increases at grade levels 7 and 9.
- The Germany and the Pacific Regions show score increases in four of the five levels tested.
- Language arts score improvement is noted across the regions for grades 7 and 11.

Mathematics

- The most significant increase in mathematics achievement is observed at grade 4.
- Increased achievement was observed for each of the five grades tested in the Mediterranean Region.
- The Panama Region shows improvement in mathematics in four of the five levels tested with dramatic increases occurring at grades 4, 7, and 9.
- Three regions: Atlantic, Pacific, and Panama show a significant increase in fourth grade mathematics.

DEMOGRAPHIC INFORMATION

Demographic information was compiled through individual student data gathered at the time of the test administration. Table 13 reports the data for DoDDS systemwide. The information included is self-explanatory; however, it may be important to note that nearly 50 percent of the student population has been exposed to the DoDDS program for *less* than 2 years.

Table 13

DoDDS System-wide Testing Program
Demographic Data

	Number of Students	Percent
Grade		
Grade 4	12,280	25.7
Grade 6	11,190	23.4
Grade 7	10,700	22.4
Grade 9	8,230	17.2
Grade 11	5,452	11.4
Region		
Atlantic	5,535	11.56
Germany	26,665	55.71
Mediterranean	4,853	10.14
Pacific	8,355	17.46
Panama	2,453	5.13
English Fluency		
Fluent in English	46,043	96.2
Not Fluent in English	1,405	2.9
Language Other Than English Used at Home		
Yes	12,907	27.0
No	33,597	70.2
Don't Know	649	1.4
Length of Time in DoDDS System		
Less Than 1 Year	10,197	21.3
1 to 2 Years	12,211	25.5
2 Years or More	24,712	51.6
Sponsor's Branch of Service		
Army	20,965	43.8
Navy	3,590	7.5
Air Force	15,785	33.0
Marines	798	1.7
U.S. Government Civilian	4,750	9.9
Non-Command Sponsored Civilian	1,466	3.1
Compensatory or Supplemental Program		
Reading	2,846	6.0
Mathematics	1,424	3.0
ESL	804	1.7

APPENDIX A
Statements of DoDDS STP Category Objectives

Grade		
Word Attack		
4	Objective 1	The student will identify a word with the same long, short, or r-controlled vowel sound as that in a given word. (Letter-Sound Correspondence)
4	Objective 2	The student will recognize a compound word and will identify the meaning of a contraction. (Structural Analysis)
4	Objective 3	The student will identify a word with the same vowel diphthong or digraph sound as that in a given word. (Phonic Analysis)
Vocabulary		
4 6 7 9 11	Objective 4	The student will identify synonyms and homonyms, use context clues, or identify the meaning of a prefix or suffix. (Word Meaning)
Reading Comprehension		
4 6 7 9 11	Objective 5	The student will demonstrate knowledge of the elements of literal comprehension. (Literal Comprehension)
4 6 7 9 11	Objective 6	The student will demonstrate knowledge of the elements of literary works. (Literary Works)
4 6 7 9 11	Objective 7	The student will demonstrate knowledge of the elements of higher-level comprehension. (Higher-Level Comprehension)
Spelling		
4 6 7 9 11	Objective 8	The student will identify those words spelled correctly in written communication. (Spelling)
Language Mechanics		
4 6	Objective 9	The student will identify the correct use of a capital letter. (Capitalization)
4 6 7 9 11	Objective 10	The student will identify the correct use of punctuation marks. (Punctuation)
4 6 7 9 11	Objective 11	The student will identify correct capitalization and punctuation in a passage, friendly letter, or business letter. (Editing Skills)

Grade		
Language Expression		
4 6 7 9 11	Objective 12	The student will identify the correct form of a word to complete a sentence. (Usage)
4 6 7 9 11	Objective 13	The student will identify the subject or verb of a sentence or will identify a sentence as being complete, incomplete, or run-on. (Sentence Development)
4 6 7 9 11	Objective 14	The student will demonstrate knowledge of the skills necessary to develop paragraphs. (Paragraph Development)
Mathematics		
4 6 7 9 11	Objective 15	The student will add using whole numbers, fractions, decimals, integers, algebraic expressions, or exponents. (Addition)
4 6 7 9 11	Objective 16	The student will subtract using whole numbers, fractions, decimals, integers, algebraic expressions, or exponents. (Subtraction)
4 6 7 9 11	Objective 17	The student will multiply using whole numbers, fractions, decimals, integers, exponents, percents, or algebraic expressions. (Multiplication)
4 6 7 9 11	Objective 18	The student will divide using whole numbers, fractions, decimals, integers, or percents. (Division)
4 6 7 9 11	Objective 19	The student will demonstrate an understanding of numeration, number sentences, and number theory. (Number Awareness)
4 6 7 9 11	Objective 20	The student will demonstrate an understanding of problem solving. (Problem Solving)
4 6 7 9 11	Objective 21	The student will demonstrate an understanding of measurement and geometry. (Measurement/Geometry)
Locational Skills		
6 7 9 11	Objective 22	The student will locate information in books. (Book Parts)
6 7 9	Objective 23	The student will locate information in the dictionary. (Dictionary Skills)
6 7 9 11	Objective 24	The student will locate information in the library. (Library Skills)
7	Objective 25	The student will locate information in consumer sources. (Consumer Skills)

APPENDIX B

Test Development

The Office of Dependents Schools (ODS) curriculum specialists for reading, language arts, and mathematics provided guidance and direction to CTB/McGraw-Hill content specialists in the selection of items to match the priority objectives. In total, 25 category objectives were derived and partitioned into the subtest areas: word attack, vocabulary, reading comprehension, spelling, language mechanics, language expression, math computation, math concepts and application, and locational skills. Each grade level tested 19-21 of these objectives. A minimum of four items were selected for each category objective. A complete description of the DoDDS category objectives is shown in Figure 2.

Utilizing the CTB/McGraw-Hill item pool, from which Form V of the Comprehensive Tests of Basic Skills (CTBS U/V) also was developed, a series of tests were constructed in the subject areas and the grade levels specified by ODS. The overriding consideration in the development of these assessment instruments was insuring the compatibility of the test items with the curricular objectives. The selection of items was based on two criteria: content and statistics. From a content point of view, the items must —

a. Satisfy the objectives.
b. Have an acceptable vocabulary and readability level.
c. Be appropriate for the target group.

From a statistical perspective, the items must —

a. Be of appropriate difficulty level.
b. Not be biased against minority groups.
c. Highly discriminate between ability levels.
d. Fit the item response theory model.
e. Have a low guessing factor.

All items received rigorous review and analyses by DoDDS curriculum coordinators and outside panels of experts.

When CTB was satisfied that a good test had been chosen, a field test edition was assembled and reviewed by DoDDS curriculum staff in Washington the first week of May.

VALIDATION OF TEST INSTRUMENTS

Field Trial

A field trial of the STP tests was conducted by CTB under ODS direction. The purpose of the field trial was to ensure that the tests would function properly. Specifically, the trial provided a means of judging the adequacy, accuracy, and appropriateness of (a) the test directions and manuals, (b) the practice test, (c) the test booklets including the sample items and the directions, (d) the time limits, (e) the difficulty level of each test as a whole, and (f) the scoring keys. The study found the tests to be soundly constructed and of high psychometric quality. Comments of the teachers who participated in the field trial were strongly supportive of the quality of tests and

the adequacy of the directions and manuals. Details regarding the field trial are provided in the "STP Field Test Report, July 1983".

External Reviews

The tests were reviewed externally by a curriculum panel and a technical advisory committee. Concurrent with the field trial, an independent review of the tests was conducted by three specialists from the Washington, D.C., public schools' Competency-Based Curriculum staff. They were:

1. Dr. Helen Turner - Reading
2. Dr. Mary White - Language Arts
3. Dr. Gordon Lewis - Mathematics

The individual reviews of these panelists showed that the tests appropriately reflected the curricular areas.

In October 1982, a meeting of the Technical Advisory Committee was convened for the purpose of reviewing all of the technical supporting material documenting the comparability of the DoDDS STP to a national norm. The committee favorably endorsed the technical procedures and the quality of the test. The Technical Advisory Committee members included:

1. Dr. Robert Linn, Chairperson, Department of Educational Psychology, University of Illinois, Urbana (Measurement and Evaluation).
2. Dr. Jason Millman, Professor, School of Education, Department of Psychology, Cornell University (Measurement and Evaluation).
3. Dr. Richard Jaeger, Professor, School of Education, University of North Carolina, Greensboro (Statistics).
4. Dr. Carol Tittle, Research Psychologist, University of North Carolina, Greensboro (Research in Tests and Measurement).
5. Dr. Gary Bitter, Professor of Education, Arizona State University, Tempe (Mathematics and Computer Education).
6. Dr. M. Trika Smith-Burke, Associate Professor, Department of Educational Psychology, New York University (Reading Education).

Publication No. 84-C-0002
March 1984